IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

In re:	Application No. 10/631,256	*Confirmation No. 1617*
Filed:	July 31, 2003	
Applicants:	Prakairut TARLTON et al.	This Appellants' Brief on Appeal was electronically filed on May 27, 2008 using EFS-Website.
Title:	SPOKEN LANGUAGE SYSTEM	
Art Unit:	2626	
Examiner:	David D. KNEPPER	
Attorney Docket:	CML01259H (90891)	
Customer No.:	22242	

Mail Stop APPEAL BRIEF -- PATENTS
Commissioner for Patents
P. O. Box 1450
Alexandria, Virginia 22313-1450

APPELLANTS' BRIEF ON APPEAL UNDER 37 C.F.R. § 41.37

Sir:

Pursuant to 37 C.F.R. § 41.37, the Applicants hereby respectfully submit the following Brief in support of their appeal.

TABLE OF CONTENTS

I. REAL PARTY IN INTEREST

The real party in interest is Motorola, Inc., a Delaware corporation having a primary place of business in Schaumburg, Illinois.

II. RELATED APPEALS AND INTERFERENCES

There are no related appeals or interferences known to appellant, the appellant's legal representative, or assignee that will directly affect, or be directly affected by or have a bearing on the Board's decision in the pending appeal.

III. STATUS OF CLAIMS

Claims 1-16 are pending and presently stand twice and finally rejected, and constitute the subject matter of this appeal.

IV. STATUS OF AMENDMENTS

No post-final amendments have been submitted.

V. SUMMARY OF CLAIMED SUBJECT MATTER

A concise explanation of this subject matter appears as follows in the form of claim subject matter maps for the independent claims (with corresponding references to the specification by page and line number (or paragraph numbering where appropriate) and to the drawing(s) (if any) by figure number and reference characters where applicable.[1]

Independent Claim 1

	Relevant Figures and Relevant Specification Pages/Line Numbers
A method for a spoken language system, comprising:	Fig. 1 and Fig. 2 Paragraph [0009]
generating a recognized sequence of words from a sequence of received spoken words;	Fig. 2 Paragraphs [0010] and [0011]
assigning a confidence score to each word in the recognized sequence of words; and	Fig. 3 Paragraphs [0011], [0013], [0018], [0026], and [0027]
adjusting nominal acoustical properties of words in a presentation of the recognized sequence of words, the adjustment performed according to the confidence score of each word.	Fig. 4, Fig. 5, and Fig. 6 Paragraphs [0011], [0013], [0015], [0016], [0017], [0026], [0027], and [0029]-[0068]

[1] There are no means plus function (or step plus function) recitations in any of the claims involved in this appeal, and therefore there is no identification of any corresponding structure, material, or acts in the specification in this regard. It will be understood that this summarization of the claimed subject matter is, in fact, a "summary" and that the Applicants do not represent or intend that this brief presentation, or the accompanying references to the drawings and the specification, comprise an exhaustive presentation in this regard. As always, the claims are to be viewed and interpreted in view of the context of the entire specification sans the Abstract.

Independent Claim 15

	Relevant Figures and Relevant Specification Pages/Line Numbers
A spoken language system, comprising:	Fig. 1 (100) and Fig. 2 (200) Paragraph [0009]
a recognition component that generates a recognized sequence of words from a sequence of received spoken words, and assigns a confidence score to each word in the recognized sequence of words; and	Fig. 1 (115 and 120), Fig. 2 (220 and 225), and Fig. 3 (310) Paragraph [0010], [0011], [0013], [0018], [0026], and [0027]
a presentation component that adjusts nominal acoustical properties of words in a presentation of the recognized sequence of words, the adjustment performed according to the confidence score of each word.	Fig. 2 (230), Fig. 4 (420), Fig. 5 (520), and Fig. 6 (620) Paragraphs [0011], [0013], [0015], [0016], [0017], [0026], [0027], and [0029]-[0068]

Independent Claim 16

	Relevant Figures and Relevant Specification Pages/Line Numbers
A portable electronic device, comprising:	
a radio transceiver that can establish a telephone call;	Fig. 1 (100) Paragraph [0012]
a recognition component that generates a recognized sequence of words from a sequence of received spoken words, and assigns a confidence score to each word in the recognized sequence of words; and	Fig. 1 (115 and 120), Fig. 2 (220 and 225), and Fig. 3 (310) Paragraph [0010], [0011], [0013], [0018], [0026], and [0027]
a presentation component that adjusts nominal acoustical properties of words in a presentation of the recognized sequence of words, the adjustment	Fig. 2 (230), Fig. 4 (420), Fig. 5 (520), and Fig. 6 (620) Paragraphs [0011], [0013],

	Relevant Figures and Relevant Specification Pages/Line Numbers
performed according to the confidence score of each word.	[0015], [0016], [0017], [0026], [0027], and [0029]-[0068]

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

Claims 1-16 are rejected under 35 U.S.C. § 103(a) given Brittan et al. (U.S. Patent No. 7,062,440 B2) ("Brittan"). Claim 7 is rejected under 35 U.S.C. § 103 given Brittan in view of Takagi et al. (U.S. Patent No. 6,205,420 B1) ("Takagi").

VII. ARGUMENT

Claims 1-6 and 8-16 were rejected under 35 U.S.C. §103(a) given Brittan. Prior to addressing the merits of the Examiner's rejection, the Applicants believe it would first be helpful to briefly describe and characterize the Brittan reference.

Brittan's FIG. 4 (reproduced below) provides a general overview of his speech system.

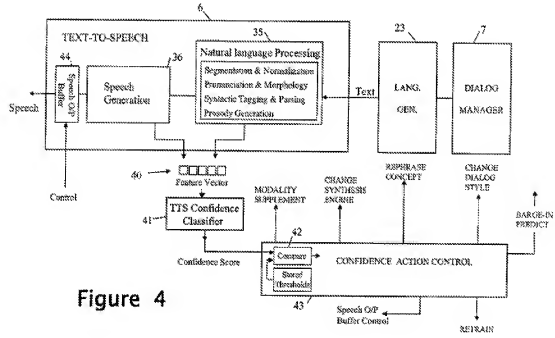

Figure 4

The Brittan application discloses a method of outputting speech through a speech synthesis engine composed of a language generator 23 and a text-to-speech converter (TTS). At each stage of the synthesis process, a TTS confidence classifier 41 can be used to modify the text to be synthesized through the assignment of a confidence score. In order to assign a confidence score, "feature values," consisting of "the degree of syntactic ambiguity in the text, the number of alternative intonation contours, and the amount of signal processing performed

in the speech generation process[2]" are used to create a "feature vector" 40. Specifically, the feature vector 40 is used by the confidence classifier 41 to incorporate the entire series of utterances as a whole[3] which serve to output a confidence score[4]. The confidence score is in turn used by the language generation block 23 or the dialogue manager 7 to modify the text that is synthesized during the process[5].

Brittan is therefore seen to teach the use of confidence scores to influence his synthesized audibilized text. Brittan, however, *only* teaches placing confidence scores broadly upon each *multi-word* general utterance or text. The Applicants disclose a method of placing a confidence score on *each and every* word within an utterance.

The Examiner admits that Brittan does not explicitly teach applying confidence scores to each word. However, the Examiner notes that Brittan teaches inserting pauses in front of certain words and suggests that adding confidence scores to each word would therefore be obvious for a person having ordinary skill in the pertinent art at the time of the invention. The Applicants respectfully disagree.

Brittan indeed teaches inserting a pause in front of certain words when there is a sufficiently low confidence score. Brittan's method of inserting a pause in front of a word, however, does *not* suggest attaching a confidence score to *each and every* word because Brittan takes each utterance *as a whole* and then outputs a confidence score based on the entire utterance. The pause, in turn, is only inserted after a confidence score is calculated from an entire utterance as a whole. Moreover, Brittan's speech system is only configured to place pauses in front of "certain words, such as non-dictionary words and other specialized terms and proper nouns[6]" or individual words that are "amenable to having a pause inserted in

[2] Brittan at column 5, lines 57-60.
[3] Brittan at column 6, lines 59-64.
[4] Brittan at column 5, lines 57-66.
[5] Brittan at column 5, line 66 to column 6, line 2.
[6] Brittan at column 7, lines 30-32.

front of them whist still sounding natural being suitable tagged[7]." It would, therefore, not be obvious to place confidence scores upon each and every word when Brittan's system already pre-determines "certain words" for which pauses can be inserted before.

Moreover, Brittan teaches away from placing a confidence score on each and every word. Several embodiments in Brittan require placing a confidence score on the entire utterance in order to carry out the principle operation of the invention. Placing the confidence score on each and every word would render the Brittan invention unsatisfactory for its intended purpose.

Brittan teaches a method of "concept rephrasing." The language generator 23 can generate a new output for the current utterance if the confidence score that is placed on the entire utterance is too low[8]. If the confidence score of the utterance is determined to be too low, the language generator 23 can be arranged to insert one or more previously determined alternative words for the phrasing for the current concept or it can simply choose to rephrase the current concept[9]. Specifically, the confidence action controller (CAC) 43 receives the speech utterance to be outputted and assigns a confidence score associated with that speech utterance. The CAC 43 compares the confidence score against a stored threshold value to determine if the speech output can be released from a buffer and outputted from the speech system. When the CAC 43 receives a low confidence score relative to the stored threshold value from a current utterance, the language generator 23 can choose to insert one or more of the stored alternative words for the current speech utterance from the speech synthesis subsystem to provide a new phrase for the current utterance with a higher confidence score[10]. Alternatively, the language generator 23 may choose to keep all the words from the current utterance and simply rearrange them in a different order to produce a new speech utterance

[7] Brittan at column 7, lines 50-53.
[8] Brittan at column 7, lines 21-25.
[9] Brittan at column 7, lines 27-34.
[10] Brittan at column 7, lines 21-29.

that has a higher confidence score[11]. If a confidence score was assigned to each and every word in the utterance, Brittan's method of "concept rephrasing" would be rendered unsatisfactory since the language generator 23 is specifically configured to recognize a confidence score that is placed on an entire utterance before it will rearrange the phrase or insert predetermined words in the utterance.

Brittan teaches a method of "dialogue style selection" using confidence scores. Brittan's FIG. 5 (reproduced below) provides a general overview of the "dialogue style selection" embodiment of his speech system.

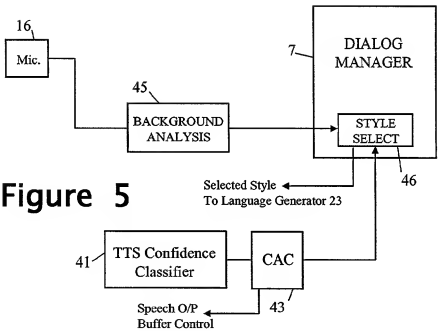

Figure 5

In an effort to provide speech in the most intelligible manner, a confidence score that is placed on an entire utterance is used to trigger a change of dialogue style. The CAC 43 is connected to a style selection block 46 of the dialog manager 7. Specifically, the dialog

[11] Brittan at column 7, line 34.

manager 7 can be pre-supplied with stored alternative scripts that correspond to a type of style for a particular utterance. If a low confidence score is assigned for a given utterance, the CAC 43 can trigger the style selection block 46 to change the style of the utterance through insertion of the stored alternative script (different words that have a similar substantive content as the utterance). If this changed style results in assigning a higher confidence score to the utterance, the CAC 43 can direct the style selection block 46 to use the newly-selected style for the subsequent utterance. Brittan's "dialogue style selection" feature would be unsatisfactory if a confidence score was placed on each and every word since the style selection block 46 is configured to select an alternative script based on receipt of a confidence score that is assigned to the entire utterance. In other words, the decision on whether to insert a stored alternative script for a low confidence utterance must be made by viewing the utterance as a whole and attempting to analyze a confidence score pertaining to each and every word would fail to suffice in adhering to the principle of this embodiment.

The Brittan reference teaches a Multi-modal output that uses the confidence score that is placed on the entire utterance. Brittan's FIG. 6 (reproduced below) provides a general overview of the "Multi-modal output" embodiment of his speech system.

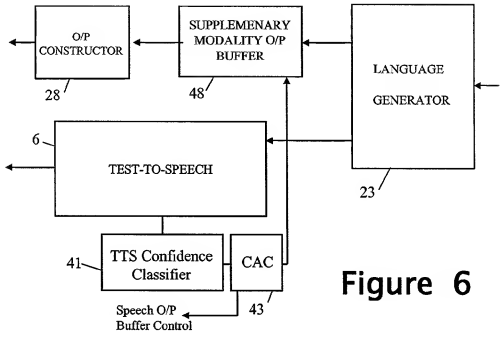

Figure 6

Brittan discloses that another output device, such as a graphical or video display, can work with the speech synthesis system. In this embodiment, the video display interacts with the speech synthesis system and the content of the video output is directly related to the confidence score that is assigned to the entire utterance. Brittan explains that the video display can be adjusted to clarify or exemplify the speech utterance[12]. Specifically, the language generator 23 is used to output speech as well as a supplementary modality output that is held in buffer 48. The supplemental modality output is only outputted when the CAC 43 places a low confidence score on an entire speech utterance. If a low confidence score is assigned to the entire speech utterance then the supplemental modality output is released from the buffer 48 and fed into the O/P Constructor 28 where it is converted to a suitable

[12] Brittan column 9, lines 22-25.

display form[13] to help clarify or exemplify the speech utterance. For example, a user may be receiving travel directions from a device both in the form of outputted speech and a map display. If the CAC 43 produces a low confidence score in relation to an utterance pertaining to a particular street name, the map display can be adjusted to provide the user with a suitable display to better allow the user to understand the travel directions. In this example, the map display could display the name of the street name in large text on the display screen[14]. Brittan's method of using a confidence score to release a supplemental modality output would be unsatisfactory if the confidence score was assigned to each and every word. Brittan's buffer 48 decides whether to output the supplemental modality and what type of display the modality should have based on the confidence score that is assigned to the entire utterance as a whole. Merely placing confidence scores on each and every word in the utterance would fail to provide a proper input as to when the supplemental modality should be released and what type of display form would be suitable for the modality, thus rendering the Brittan system unsatisfactory.

The Brittan reference teaches a "barge-in prediction" embodiment that uses the confidence score of an entire utterance to assess whether a user is likely to barge in to correct an error in the entire utterance. Brittan's FIG. 8 (reproduced below) provides a general overview of the "barge-in prediction" embodiment of his speech system.

[13] Brittan at column 9, lines 30-36.
[14] Brittan at column 9, lines 15-20.

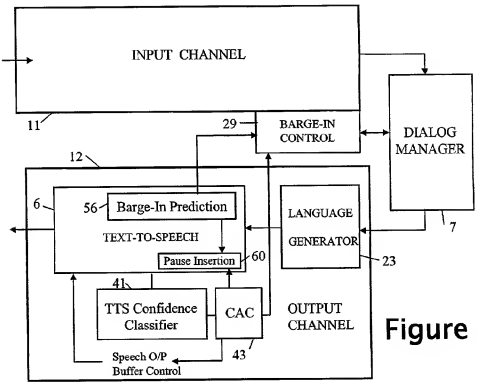

Figure 8

Brittan's barge-in feature differs considerably from the Applicants' disclosure. Brittan describes a barge-in feature that requires a confidence score to be placed on the entire utterance in order to function in a satisfactory manner. The barge-in feature allows the user to easily correct a possible error in the utterance if a low confidence score is assigned to the entire utterance. Specifically, if the CAC 43 receives an utterance that has a low confidence score, the CAC 43 will indicate the low score to the barge-in control 29. The barge-in control will cause the system to become increasingly sensitive and more alert to accept a barge-in correction from a user[15]. This feature would fail to work properly if a confidence score was assigned to each and every word. Under the Brittan system, the barge-in control 29 only makes the

[15] Brittan at column 10, lines 36-42.

system ready for a barge-in when a low confidence score is placed upon the entire utterance. Assigning a confidence score to each and every word in the utterance would likely provide varying scores to the barge-in control 29 and would fail to provide the barge-in control 29 with proper guidance for when to prepare the system for a possible barge-in from the user.

Thus, Brittan's teaching of placing a confidence score on a series of words or phrases does not teach or suggest the Applicants' use of a confidence score. In substance the two approaches differ considerably. Brittan teaches, and *only* teaches, the determination of a confidence rating as a function of a group of words whereas the Applicants specifically set forth and require placing a confidence score on *each and every* word. Not only does Brittan teach placing a confidence score on an entire utterance but failure to place the confidence score on the entire utterance would cause the invention to be made unsatisfactory for many of its intended purposes. The Applicants respectfully traverse these rejections and respectfully request reconsideration.

Claims 1, 15 and 16

Independent claims 1, 15 and 16 are not obvious in view of the argument above. Claims 1, 15 and 16 all recite the language of "assigning a confidence score to each word in the recognized sequence of words." This limitation is not obvious in view of the argument above. The Applicants therefore respectfully submit that Claim 1, 15 and 16 is not obvious for a variety of reasons, including those expressed above.

Claims 2-14

Claims 2-14 are ultimately dependent upon claim 1, which claim has been shown to be allowable above. While the Applicants believe that other arguments are available to highlight the allowable subject matter presented in various of these remaining dependent claims, the Applicants also believe that the comments set forth herein regarding allowability of the

independent claims are sufficiently compelling to warrant present exclusion of such additional points for the sake of brevity and expedited consideration

VIII. CLAIMS APPENDIX

1. (Original) A method for a spoken language system, comprising: generating a recognized sequence of words from a sequence of received spoken words; assigning a confidence score to each word in the recognized sequence of words; and adjusting nominal acoustical properties of words in a presentation of the recognized sequence of words, the adjustment performed according to the confidence score of each word.

2. (Original) The method according to claim 1, wherein adjusting comprises: adjusting the presentation using a lengthened interword pause proximate to a word having a low confidence score, wherein the lengthened interword pause is recognizably greater than interword pauses otherwise used for words having a confidence score within a normal range.

3. (Original) The method according to claim 2, wherein the lengthened interword pause is inserted directly following the word having a low confidence score.

4. (Original) The method according to claim 2, wherein the lengthened interword pause is inserted after a group of words that includes the word having a low confidence score.

5. (Original) The method according to claim 2, wherein the lengthened interword pause is inserted following the word having a low confidence score, and the duration of the pause is determined based on an amount by which the confidence score indicates a confidence below the normal range.

6. (Original) The method according to claim 2, wherein the lengthened interword pause is inserted following the word having a below normal confidence score, and a duration of the lengthened interword pause is determined based on a likely duration of the corrective response.

7. (Original) The method according to claim 6, wherein the likely duration of the corrective response is one of a duration of a button press and a duration of the words predicted to be spoken during the lengthened interword pause.

8. (Original) The method according to claim 2, wherein the lengthened interword pause is inserted directly preceding the word having a below normal confidence score.

9. (Original) The method according to claim 8, wherein the duration of the lengthened interword pause is increased for lower confidence scores.

10. (Original) The method according to claim 1, wherein adjusting comprises: modifying a nominal value of one or more of a set of acoustical features for a word having a confidence score outside of a normal range.

11. (Original) The method according to claim 10, wherein the set of acoustical features comprises interword pause, duration, pitch range, intonational contour, intensity, phonation type, and precision of articulation.

12. (Original) The method according to claim 10, wherein the modifying comprises at least one of: increasing at least one of the interword pause, the duration of the word, the pitch range of the word, the loudness of the word, and the precision of articulation of the word when the confidence score indicates a lower than nominal confidence; and decreasing at least

one of the interword pause, the duration of the word, the pitch range of the word, the loudness of the word, and the precision of articulation of the word when the confidence score indicates a higher than nominal confidence.

13. (Original) The method according to claim 10, wherein the set of acoustical features further comprises a duration change of each syllable of the word, and wherein a differential change of the duration of each syllable is determined by a lexical stress parameter of the syllable.

14. (Original) The method according to claim 10, wherein adjusting comprises: adjusting the presentation using a phrase contour that conveys uncertainty within a group of words that includes a word having a confidence score below the normal range.

15. (Original) A spoken language system, comprising: a recognition component that generates a recognized sequence of words from a sequence of received spoken words, and assigns a confidence score to each word in the recognized sequence of words; and a presentation component that adjusts nominal acoustical properties of words in a presentation of the recognized sequence of words, the adjustment performed according to the confidence score of each word.

16. (Original) A portable electronic device, comprising: a radio transceiver that can establish a telephone call; a recognition component that generates a recognized sequence of words from a sequence of received spoken words, and assigns a confidence score to each word in the recognized sequence of words; and a presentation component that adjusts nominal acoustical properties of words in a presentation of the recognized sequence of words, the adjustment performed according to the confidence score of each word.

IX. EVIDENCE APPENDIX

None.

X. RELATED PROCEEDINGS APPENDIX

None.

XI. CONCLUSION

In view of the foregoing discussion, the Applicants respectfully request reversal of the rejected, pending claims.

Respectfully submitted,

FITCH, EVEN, TABIN & FLANNERY

Dated: 5/27/2008

Steven G. Parmelee
Registration No. 28,790

120 South LaSalle Street, Suite 1600
Chicago, Illinois 60603-3406
Telephone (312) 577-7000
Facsimile (312) 577-7007
505519